Crystal structure of HINT from *Helicobacter pylori*

K. F. Tarique, S. Devi, S. A. Abdul Rehman and S. Gourinath*

School of Life Science, Jawaharlal Nehru University, New Delhi, Delhi 110 067, India. *Correspondence e-mail: sgourinath@mail.jnu.ac.in

Received 6 October 2015

Accepted 4 December 2015

Edited by L. J. Beamer, University of Missouri,
USA**Keywords:** histidine trinucleotide repeat;
Helicobacter pylori; protein kinase C-interacting
protein; cell cycle; crystal structure.**PDB reference:** HINT, 4zg1**Supporting information:** this article has
supporting information at journals.iucr.org/f

Proteins belonging to the histidine triad (HIT) superfamily bind nucleotides and use the histidine triad motif to carry out dinucleotidyl hydrolase, nucleotidyl-transferase and phosphoramidite hydrolase activities. Five different branches of this superfamily are known to exist. Defects in these proteins in humans are linked to many diseases such as ataxia, diseases of RNA metabolism and cell-cycle regulation, and various types of cancer. The histidine triad nucleotide protein (HINT) is nearly identical to proteins that have been classified as protein kinase C-interacting proteins (PKCIs), which also have the ability to bind and inhibit protein kinase C. The structure of HINT, which exists as a homodimer, is highly conserved from humans to bacteria and shares homology with the product of fragile histidine triad protein (FHIT), a tumour suppressor gene of this superfamily. Here, the structure of HINT from *Helicobacter pylori* (HpHINT) in complex with AMP is reported at a resolution of 3 Å. The final model has R and R_{free} values of 26 and 28%, respectively, with good electron density. Structural comparison with previously reported homologues and phylogenetic analysis shows *H. pylori* HINT to be the smallest among them, and suggests that it branched out separately during the course of evolution. Overall, this structure has contributed to a better understanding of this protein across the animal kingdom.

1. Introduction

The histidine triad (HIT) families of proteins are hydrolases that act on the α -phosphates of ribonucleotides. They are ubiquitously present throughout the kingdoms of life from prokaryotes and archaea to eukaryotes. These proteins have a conserved HXHXHXX motif (where X denotes a hydrophobic residue) that is a signature sequence of this family (Krakowiak & Fryc, 2012; Martin *et al.*, 2011; Brenner, 2002; Lima *et al.*, 1997). The HIT family has been classified into five major branches on the basis of sequence, domain structure, substrate specificity, catalytic mechanism and biological function (Martin *et al.*, 2011). These are galactose-1-phosphate uridylyltransferase (GalT; Brenner, 2002), aprataxin (Kijas *et al.*, 2006), scavenger decapping enzyme (Dcps; Liu *et al.*, 2002), fragile histidine triad protein (FHIT; Huang *et al.*, 2004) and histidine triad nucleotide-binding protein (HINT; Krakowiak & Fryc, 2012; Martin *et al.*, 2011; Ozga *et al.*, 2010). Disruption of the proper functioning of these proteins causes severe diseases such as galactosaemia, oculomotor apraxia 1 (a neurodegenerative disease), defects in RNA metabolism and various types of cancer (Martin *et al.*, 2011).

HINT catalyzes the hydrolysis of the adenine base in a number of nucleoside phosphoramidites, resulting in the production of monophosphate nucleotides (Lima *et al.*, 1997). Protein kinase C-interacting protein (PKCI) is a member of the HIT family that is nearly identical to HINT but has a distinct role in binding and inhibiting protein kinase C (Lima *et al.*, 1996; Klein *et al.*, 1998; Brenner *et al.*, 1999). PKCI and

HINT are normally found as dimers, and their monomer forms an $\alpha+\beta$ meander, which is a type of $\alpha+\beta$ fold (Maize *et al.*, 2013). More than 20 HINT structures have currently been deposited in the Protein Data Bank (PDB) and are derived from a wide variety of organisms, with some of these HINT structures known to function as protein kinase C inhibitors. The monomer of HINT specifically consists of five antiparallel β -sheets (denoted as B1–B5) and two α -helices (H1–H2) that interact extensively in a symmetrical fashion with another monomer to form the dimer (Lima *et al.*, 1997). These 20 HINT structures are very similar overall, but their C-termini display different lengths and orientations, which is important since the C-terminus is assumed to determine substrate specificity and biological function in these proteins (Lima *et al.*, 1997). They share a common catalytic mechanism and possess conserved histidyl residues in the active site; catalysis proceeds through the formation of a covalent nucleotidyl phosphohistidyl intermediate (Lima *et al.*, 1997). Many of the HINT-family homologues have a conserved zinc-binding motif C-X-X-C (where X is a hydrophobic residue) in the α 1-helix region of the tertiary structure (Klein *et al.*, 1998). Zinc is coordinated by these cysteine residues as well as the first histidine residue of the conserved triad motif (Lima *et al.*, 1996). Structural and functional studies of human and rabbit HINT1 revealed that this protein is involved in regulation of transcription and is a potential tumour suppressor controlling regulation of the cell cycle (Maize *et al.*, 2013; Dolot *et al.*, 2011). While HINT and PKCI are functionally different, many key structural features of HINT have been found to be very similar to those of PKCI, pointing towards a common evolutionary origin for both.

The fundamental roles of HINT and PKCI in biological function combined with their ubiquitous presence throughout the animal kingdom, including in pathogenic bacteria, make them attractive targets for drugs to combat diseases caused by such bacteria. *Helicobacter pylori* is a Gram-negative pathogenic bacterium that causes atrophic gastritis upon infection, which may progress to gastric cancer if left untreated. Gastric cancer remains one of the most common cancers, and is the second leading cause of cancer-related deaths worldwide (Parkin *et al.*, 1999).

Here, we report the crystal structure of a HINT protein (also annotated as protein kinase C inhibitor) from *H. pylori* (HpHINT) in complex with AMP at a resolution of 3 Å. This work is part of our overall aim to determine the structures of novel drug targets from *H. pylori*, an effort similar to the ongoing research towards increasing the structural information available for *Entamoeba histolytica* and *Mycobacterium tuberculosis* drug discovery (Baugh *et al.*, 2015; Lorimer *et al.*, 2015).

2. Materials and methods

2.1. Macromolecule production

Gene ID HP0404, annotated as a protein kinase C inhibitor or an ADP hydrolase from the HINT protein family

Table 1

Macromolecule-production information.

Source organism	<i>H. pylori</i> strain 26695
DNA source	Genomic DNA
Forward primer	5'-CTAGCTAGCATGAATGTGTTTGAATAA-TC-3'
Reverse primer	5'-CGGCTCGAGATGTTTGTCTCCGCTTAAAA-3'
Cloning vector	pET-21c
Expression vector	pET-21c
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced	MNVFEKIIQGEIPCSKILENERFLSFYDINPKAK-VHALVIPKQSIQDFNGITPELMAQMTSIFIEV-VEKLGIKEKGYKLLTNVGKNAGQEVMLHFHILSGDKHLEHHHHHH

(HpHINT), was amplified by polymerase chain reaction (PCR). The amplified product of HpHINT (315 bp) was digested with *Nhe*I and *Xho*I and ligated into an expression vector with a C-terminal His tag (Table 1).

2.2. Expression and purification of HpHINT in *Escherichia coli*

The pET-21c-HpHINT plasmid was transformed into a bacterial expression system for recombinant protein production. The positive colonies were picked from LB plates containing ampicillin and inoculated into 50 ml LB medium for overnight growth. 1% of this culture was inoculated into a secondary culture containing 50 mg l⁻¹ ampicillin. The culture was incubated at 37°C until the OD₆₀₀ reached 0.6–0.8. Protein expression was initiated by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 250 μ M. The temperature was lowered to 30°C after induction and the culture was allowed to grow for a further 5 h. The cells were harvested at 4°C by centrifugation at 8000 rev min⁻¹ for 5 min and the pellet was suspended (at 1 g of pellet per 10 ml) in buffer A consisting of 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM phenylmethanesulfonyl fluoride (PMSF), 0.1% Triton X-100, 0.2 mM EDTA, 10 mM imidazole, 0.05 mg ml⁻¹ lysozyme, 5 mM β -mercaptoethanol (β Me). Lysed cells were sonicated and then centrifuged at 18 000 rev min⁻¹ for 1 h. The cleared supernatant was applied onto an Ni Sepharose column and the column was equilibrated with 50 ml buffer B (50 mM Tris pH 7.5, 150 mM NaCl, 5 mM β Me, 20 mM imidazole). The protein was eluted in buffer C consisting of 50 mM Tris pH 7.5, 150 mM NaCl, 400 mM imidazole, 5% glycerol, 5 mM β Me.

For further purification and the removal of imidazole, the protein was concentrated and loaded onto a HiLoad 16/60 Superdex 200 column pre-equilibrated with buffer D (10 mM Tris pH 7.5, 150 mM NaCl, 5% glycerol, 5 mM β Me). The protein was eluted at a flow rate of 1 ml min⁻¹. The purity of the desired protein was checked using 12% SDS-PAGE and the purified fractions were pooled and concentrated using a Centricon concentrating device (Millipore) with a molecular-weight cutoff of 3 kDa. Glycerol was added to increase the solubility and stability of the protein.

The identity of the HpHINT protein was confirmed by tryptic digestion and mass fingerprinting using a Bruker electrospray ionization-ion-trap mass spectrometer at the

Advanced Instrumentation Research Facility, Jawaharlal Nehru University. The protein concentration was determined by measuring the UV absorption at 280 nm. An approximate extinction coefficient of $2980\text{ M}^{-1}\text{ cm}^{-1}$ was calculated using the *ExPASy* server.

2.3. Crystallization

The purified HpHINT protein was concentrated and incubated with 3 mM adenosine monophosphate (AMP) for 2–3 h prior to crystallization. After rigorous optimization using the hanging-drop vapour-diffusion method, needle-shaped crystals appeared in 3–4 d using Morpheus crystallization screen condition No. H4 (Gorrec, 2009; Table 2).

2.4. Data collection and processing

The needles were mounted in cryoloops and flash-cooled in liquid nitrogen. Data were collected from these crystals at the European Synchrotron Research Facility (ESRF), Grenoble, France (Table 3). The data were indexed and processed using *iMosflm* (Battye *et al.*, 2011) and scaled with *SCALA* (Evans,

Table 2
Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	24-well plate
Temperature (K)	277
Protein concentration (mg ml ⁻¹)	15
Buffer composition of protein solution	10 mM Tris pH 8.0, 150 mM NaCl, 5% glycerol, 5 mM β Me
Composition of reservoir solution	12.5% PEG 3350, 12.5% PEG 1000, 12.5% MPD, 0.03 M amino-acid mix
Volume and ratio of drop	2 μ l, 1:1
Volume of reservoir (ml)	0.5

2006) from the *CCP4* suite (Winn *et al.*, 2011). These crystals were shown to belong to an orthorhombic space group with ten HpHINT protomers per asymmetric unit.

2.5. Structure solution and refinement

The HpHINT structure was determined by molecular replacement using the structure of HINT from *Coxiella thermocellum* (CtHINT; PDB entry 1xqu, 46% sequence identity; Southeast Collaboratory for Structural Genomics,

Figure 1

Structure of HpHINT. (a) Ribbon diagram of the HpHINT dimer, along with a stick diagram of bound AMP. The dimer has a conical shape. The structure consists mainly of antiparallel β -strands (denoted as B1–B5) surrounded by α -helices (H1 and H2) and short 3_{10} -helices (G1 and G2) as well as loops (L1–L6). The dimer is spanned by one ten-stranded antiparallel sheet also known as an $\alpha\beta$ -meander. Secondary structures leading to the formation of the active-site pocket (H1, L1, L3, L6, B2, B3 and B5) are also shown. The left inset shows an expanded view of the bound AMP along with the corresponding $2F_o - F_c$ electron density at a 1σ cutoff. (The active site to which this AMP bound is shown as faded ribbons.) The right inset shows an expanded view of a stick diagram of a portion of strand B5 and loop L6 and the corresponding $1\sigma\ 2F_o - F_c$ density. (b) Topology diagram of an HpHINT monomer prepared using *TopDraw* (Bond, 2003). (c) Surface-topology image of the HpHINT active-site pocket that harbours the bound AMP, which is also shown as a stick image.

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Beamline BM14, ESRF
Wavelength (Å)	0.95372
Temperature (K)	100
Detector	MAR scanner 345 mm plate CCD
Crystal-to-detector distance (mm)	275.66
Rotation range per image (°)	1
Total rotation range (°)	200
Exposure time per image (s)	2
Space group	$P2_12_12_1$
a, b, c (Å)	48.90, 75.64, 335.09
α, β, γ (°)	90, 90, 90
Mosaicity (°)	0.3
Resolution range (Å)	41.07–2.95 (3.1–2.95)
Total No. of reflections	157391
No. of unique reflections	25808
Completeness (%)	96.1 (96.6)
Multiplicity	6.1 (6.3)
$\langle I/\sigma(I) \rangle$	7.8 (2.2)
$R_{\text{r.i.m.}}$	0.2 (0.9)
Overall B factor from Wilson plot (Å ²)	51

Table 4

Structure solution and refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	41.07–2.95 (3.027–2.950)
Completeness (%)	94.5
No. of reflections, working set	24499 (1819)
No. of reflections, test set	1298 (93)
Final R_{cryst}	0.265 (0.364)
Final R_{free}	0.282 (0.385)
No. of non-H atoms	
Protein	7933
Ligand	115
Water	0
Total	8048
R.m.s. deviations	
Bonds (Å)	0.009
Angles (°)	1.485
Average B factors (Å ²)	
Protein	57.3
Ligand	68.1
Ramachandran plot	
Most favoured (%)	97.8
Allowed (%)	1.6

unpublished work). This CtHINT structure and the aligned sequences of HpHINT and CtHINT were input into CHAINSAW (Schwarzenbacher *et al.*, 2004) to produce the specific search model used for molecular replacement, which was then carried out using Phaser (McCoy *et al.*, 2007) from the CCP4 suite (Winn *et al.*, 2011). The best solution with a high translation-function Z-score (TFZ score) and log-likelihood gain (LLG) was selected. An initial cycle of restrained refinement using this solution and REFMAC5 (Murshudov *et al.*, 2011) resulted in a decrease in the R factor to 39%. The remaining parts of the polypeptide were built manually with Coot (Emsley *et al.*, 2010) and AMP molecules were added according to the difference Fourier map. Multiple rounds of manual adjustments of the model in Coot and refinement with REFMAC5 were carried out. The final model of HpHINT consists of 7933 protein atoms and five AMP molecules, fitted well into the electron density and yielded good refinement statistics with an R and R_{free} of 26 and 28%, respectively (Table 4). The structure factors and coordinates of this model have been deposited in the Protein Data Bank (PDB) as entry 4zgl, and their quality was checked using PROCHECK (Laskowski *et al.*, 1993).

3. Results and discussion

3.1. Overview of the crystal structure of HpHINT

The crystal structure of native HpHINT, determined to a resolution of 2.95 Å, was refined to an R and R_{free} of 26 and 28%, respectively. Fig. 1(a) shows the quality of the electron density for the AMP molecule and one of the representative regions of the protein.

Analysis of the crystallographic data revealed five conical shaped HpHINT homodimers (*i.e.* ten monomers) per asymmetric unit. This oligomeric state is consistent with gel-filtration experiments, where the functional form of HpHINT was found to be a dimer (data not shown). The total surface

Figure 2

Dimeric interface of HpHINT. Schematic representation of important intersubunit polar and hydrophobic interactions between various amino-acid residues in strand B4 of each HpHINT monomer.

area buried by the interactions between the two monomers was calculated to be approximately 1100 Å² using the *PISA*

server (Krissinel & Henrick, 2007), with these interactions being mostly hydrophobic in nature.

HpHINT belongs to the α/β class of proteins consisting mainly of antiparallel β -sheets with segregated α and β regions. The HpHINT monomer consists of two α -helices (denoted as H1 and H2), five β -strands (B1–B5), two short 3_{10} -helices (G1 and G2) and six named loops (L1–L6) (Figs. 1*a* and 1*b*). Residues 1–71 in the N-terminal region constitute a typical HINT family-like protein fold that consists of a three-layer $\alpha/\beta/\alpha$ fold. Residues 76–102 constitute the C-terminus.

Despite the presence of two active sites per HpHINT homodimer, only one ligand was observed to bind to the homodimer according to the electron density. Yet, in the current crystal structure, no significant change in the structure or size of the active site was found between the two protomers (r.m.s.d. of 0.3 Å). Note also that the homodimer formed with extensive hydrophobic contacts and polar interactions (protein–protein main chain–main chain hydrogen bonds) between the B4 strands of the two monomers (Fig. 2), resulting in the formation of a ten-stranded antiparallel sheet, classified as an $\alpha\beta$ -meander, that spans the homodimer.

3.2. Geometry of the active site and comparison of HpHINT with other homologous enzymes

Inspection of the crystal structure revealed that each of the two active sites of the HpHINT homodimer is hydrophobic in nature and consists of the H1 helix, the L1, L3 and L6 loops and a floor consisting of the B1, B2, B3 and B5 antiparallel strands to form a pocket that was observed to contain the bound adenosine monophosphate (AMP) in one of the active sites (Fig. 1*c*). The area and volume of the active-site pocket were determined using the *CASTp* server (Dundas *et al.*, 2006)

Figure 3
Interactions of HpHINT with AMP. The interactions of the HpHINT protein with bound AMP at the active site. The plots were generated by *LigPlot+* (Wallace *et al.*, 1995). Hydrogen bonds are shown as green dotted lines, while arcs represent residues making nonbonded contacts with the ligands.

Figure 4
Structural superposition. (*a*) Structural superposition of HpHINT (green) on homologues from *E. coli* (magenta), rabbit (cyan) and *Homo sapiens* (yellow) reveal similar core structures but different lengths for the corresponding (superposed) secondary-structural elements. (*b*) Superposition of the active sites of these structures shows that the amino-acid residue coordination and the specific orientations of the residues required for the binding of the nucleotides (AMP/GMP) are conserved and superpose well with each other.

and were found to be about $147 \text{ \AA}^2$ and $248 \text{ \AA}^3$, respectively. The amino-acid residues interacting with AMP at the active site and their respective interatomic distances in the HpHINT crystal structure are sketched in Fig. 3. AMP was observed to make numerous stabilizing interactions with the protein, with the ribose sugar interacting with Asp28 and the 5'-phosphate group of AMP interacting with Asn82, Gly88, Glu90, Val91, His95 and His97. The latter two His residues appear to be in an optimal orientation for an in-line attack of the α -phosphate position to lead to the formation of a nucleotidyl phosphotidyl intermediate before catalysis.

These amino-acid residues and the associated hydrogen bonds and hydrophobic interactions at the active site were found to be largely conserved or similar to other homologous HINT structures (Fig. 4). Homologous X-ray crystal structures of HINT enzymes from other organisms are listed in Supplementary Fig. S1 and Fig. 5. They all share a typical crescent-shaped, β -sheet structure, a similar fold and the characteristic C-terminal active-site motif HXHXHXX, where *X* is a hydrophobic residue. Pairwise alignments and three-dimensional structural superpositions of HpHINT with these homologous proteins were carried out with the DALI (Holm & Sander, 1995) and MAPSCI (Ilinkin *et al.*, 2010) web-based servers to determine structural differences and to identify the common core secondary-structural elements. The lengths of these enzymes vary from 101 to 157 amino-acid residues, but were found using these servers to have a common core size of about

81 residues with an r.m.s.d. value of $0.7 \text{ \AA}$. Unlike many other histidine triad nucleotide-binding proteins, HpHINT does not possess the conserved cysteine residues [C-X-(I/V)-F-C] for binding zinc, and any attempt to co-crystallize zinc with HpHINT resulted in protein aggregation. Besides this, HpHINT showed an absolute requirement for AMP for crystallization.

Substantial structural differences were found between HpHINT and the other homologous structures. These differences are highlighted in terms of r.m.s.d.s, the number of C^α atoms aligned and sequence identity in Supplementary Table S1. The structural comparisons revealed a common folding pattern among HINT homologues, with the majority of the structural differences occurring in the C-terminal region (Fig. 5). These variations in the homologous HINT enzymes are owing to differences in the length, size, amino-acid residue composition and orientation of the C-terminal regions, which appear to make this region flexible in HINT-subfamily proteins. These structural variations have previously been suggested to account for the great variability in substrate specificity across the members of this superfamily (Lima *et al.*, 1997), but conclusive evidence for this hypothesis has not yet been obtained. These differences are also reflected in the structure-based phylogenetic tree constructed using the POSA server (Li *et al.*, 2014), which showed a common origin for all of the HINT homologues (chosen for the present study), which thus formed a single clade (Fig. 6).

Figure 5

HINT homologues. Note the various orientations of the COOH-terminal domains (magenta) in the HINT homologues. The PDB codes for these structures and the organisms to which they belong are listed in Fig. 6.

Figure 6

Structure-based phylogenetic tree. Representative structures of various HINT homologues were obtained from the PDB and a phylogenetic tree was constructed. The tree shows a common origin for these proteins. The PDB entries used for the structural comparison are indicated in parentheses.

4. Conclusions

The crystal structure of the putative protein kinase C inhibitor HphINT from *H. pylori* in complex with AMP was obtained at a resolution of 3 Å. It is a dimeric protein containing conserved mononucleotide-binding sites in each subunit, but in the present structure AMP was observed to be bound to the active site of only one subunit; however, no major structural change was observed in the geometry of the apo partner. The overall β -meander structure of HphINT is similar to that of homologous HINT structures, but considerable differences in the length and the orientation of the COOH-terminal region were observed between these homologues. Strangely, protein kinase C homologues are absent from bacteria, archaea and protista, yet the genomic databases of these organisms appear to have incorrectly annotated the HINT protein as PKCI. While HINT and PKCI are structurally very similar, HINT is not a PKCI as they have different functions. In the present case of HphINT the absence of a zinc-binding motif, the presence of AMP at the active site and the occurrence of a typical β -meander structure suggest the protein to be a HINT, but a detailed functional characterization is required to determine the exact role of this protein *in vivo*.

5. Related literature

The following references are cited in the Supporting Information for this article: Gouet *et al.* (2003) and McWilliam *et al.* (2013).

Acknowledgements

We thank the Department of Biotechnology, Government of India for funding. KFT, SD and SAAR thank DBT, ICMR and CSIR for fellowships. We also thank the staff of BM14, ESRF for helping us in the collection of high-resolution X-ray data.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Baugh, L. *et al.* (2015). *Tuberculosis*, **95**, 142–148.
- Bond, C. S. (2003). *Bioinformatics*, **19**, 311–312.
- Brenner, C. (2002). *Biochemistry*, **41**, 9003–9014.
- Brenner, C., Bieganski, P., Pace, H. C. & Huebner, K. (1999). *J. Cell. Physiol.* **181**, 179–187.
- Dolot, R., Ozga, M., Krakowiak, A. & Nawrot, B. (2011). *Acta Cryst.* **D67**, 601–607.
- Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). *Nucleic Acids Res.* **34**, W116–W118.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Gorrec, F. (2009). *J. Appl. Cryst.* **42**, 1035–1042.
- Gouet, P., Robert, X. & Courcelle, E. (2003). *Nucleic Acids Res.* **31**, 3320–3323.
- Holm, L. & Sander, C. (1995). *Trends Biochem. Sci.* **20**, 478–480.
- Huang, K., Arabshahi, A., Wei, Y. & Frey, P. A. (2004). *Biochemistry*, **43**, 7637–7642.
- Ilinkin, I., Ye, J. & Janardan, R. (2010). *BMC Bioinformatics*, **11**, 71.
- Kijas, A. W., Harris, J. L., Harris, J. M. & Lavin, M. F. (2006). *J. Biol. Chem.* **281**, 13939–13948.
- Klein, M. G., Yao, Y., Slosberg, E. D., Lima, C. D., Doki, Y. & Weinstein, I. B. (1998). *Exp. Cell Res.* **244**, 26–32.
- Krakowiak, A. & Fryc, I. (2012). *Postepy Biochem.* **58**, 302–313.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.
- Li, Z., Natarajan, P., Ye, Y., Hrabe, T. & Godzik, A. (2014). *Nucleic Acids Res.* **42**, W240–W245.
- Lima, C. D., Klein, M. G. & Hendrickson, W. A. (1997). *Science*, **278**, 286–290.
- Lima, C. D., Klein, M. G., Weinstein, I. B. & Hendrickson, W. A. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 5357–5362.
- Liu, H., Rodgers, N. D., Jiao, X. & Kiledjian, M. (2002). *EMBO J.* **21**, 4699–4708.
- Lorimer, D. D., Choi, R., Abramov, A., Nakazawa Hewitt, S., Gardberg, A. S., Van Voorhis, W. C., Staker, B. L., Myler, P. J. & Edwards, T. E. (2015). *Acta Cryst.* **F71**, 572–576.
- Maize, K. M., Wagner, C. R. & Finzel, B. C. (2013). *FEBS J.* **280**, 3389–3398.
- Martin, J., St-Pierre, M. V. & Dufour, J.-F. (2011). *Biochim. Biophys. Acta*, **1807**, 626–632.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. (2013). *Nucleic Acids Res.* **41**, W597–W600.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Ozga, M., Dolot, R., Janicka, M., Kaczmarek, R. & Krakowiak, A. (2010). *J. Biol. Chem.* **285**, 40809–40818.
- Parkin, D. M., Pisani, P. & Ferlay, J. (1999). *CA Cancer J. Clin.* **49**, 33–64.
- Schwarzenbacher, R., Godzik, A., Grzechnik, S. K. & Jaroszewski, L. (2004). *Acta Cryst.* **D60**, 1229–1236.
- Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). *Protein Eng. Des. Sel.* **8**, 127–134.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.